DOCUMENT RESUME

ED 228 280 TM 830 156

AUTHOR Martin, Paula H.
TITLE Meta-Analysis, Meta-Evaluation and Secondary Analysis.
PUB DATE Oct 82
NOTE 37p.
PUB TYPE Information Analyses (070)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Bilingual Education Programs; *Educational Research; Elementary Secondary Education; Federal Programs; *Program Evaluation; *Research Methodology; *Statistical Analysis; *Statistical Data
IDENTIFIERS Elementary Secondary Education Act Title VII; Glass (G V); *Meta Analysis; *Meta Evaluation; Secondary Analysis

ABSTRACT

Meta-analysis, meta-evaluation and secondary analysis are methods of summarizing, examining, evaluating or re-analyzing data in research and evaluation efforts. Meta-analysis involves the summarization of research findings to come to some general conclusions regarding effects or outcomes of a given treatment/project/program. Glass's approach standardizes various effect measures and controls for these in analyzing data. Meta-evaluation is a method of evaluation research examining evaluation methodologies, procedures, data analysis techniques, interpretation of results, and the validity and reliability of conclusions. Secondary analysis, as defined by Glass, is, "the re-analysis of data for the purpose of answering the original research question with better statistical techniques or answering new questions with old data." A review of the literature related to these methodologies gives examples of actual studies using these techniques. Specifics on meta-evaluation in federally funded bilingual education programs illustrate the methodology. As current budgetary cutbacks affect state and federal programs, meta-analysis and meta-evaluation are assuming important roles. (CM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

P. H. Martin

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED228280

META-ANALYSIS, META-EVALUATION

AND

SECONDARY ANALYSIS

by

Paula H. Martin

October 1982

Paula H. Martin, Ed.D.
California Evaluation Services
2090 Green Street, No. 27
San Francisco, CA 94123

TM 830 156

*This paper was originally written as part of an independent study for Dr. Michael Scriven, University of San Francisco, Spring 1981.

Paula Martin, Ed.D.
California Evaluation Services
2090 Green Street, No. 27
San Francisco, CA 94123

"Meta-Analysis, Meta-Evaluation and Secondary Analysis"

This paper presents an overview of meta-analysis, meta-evaluation and secondary analysis as methods of summarizing, examining, evaluating or re-analyzing data in research and evaluation efforts.

A review of the literature related to these methodologies is included and examples of actual studies where these techniques were implemented.

The paper includes some specifics on meta-evaluation in federally-funded bilingual education programs which was the technique employed by the author in her doctoral dissertation.

Conclusions are drawn as to the relevancy of these methodologies in improving existing and future research and evaluation efforts.

TABLE OF CONTENTS

I. Meta-Analysis.....1-8
Meta-Analysis Technique.....2-3
Meta-Analysis: Review of the Literature and Relevant Studies.....3-8

II. Meta-Evaluation.....8-19
Meta-Evaluation: Review of the Literature and Relevant Studies 8-11
Meta-Evaluation and Bilingual Education.....12
CES Suggested Criteria for Bilingual Program Evaluations 14-16

III. Secondary Analysis.....20-21

IV. Conclusions 21-23

LIST OF TABLES

Table I Cook and Gruders' "Models of Meta-Evaluation," (1978) 11

Table II Data Analysis Section:"CES Meta Evaluation Instruments ".....14
Martin (1980)

Table III CES Meta-Evaluation Instruments-Martin(1980).....15

Table IV CES Supplemental Data Sheet-Meta-Evaluation Instruments.....16
Martin(1980)

META-ANALYSIS, META-EVALUATION AND SECONDARY ANALYSIS

This paper will discuss meta-analysis, meta-evaluation and secondary analysis as methods of examining and summarizing research and evaluation findings.

A review of the literature pertaining to these topics will be included and a discussion of meta-analysis and meta-evaluation efforts as they pertain to federally-funded compensatory education programs such as ESEA Title I and ESEA Title VII (Bilingual Education).

I. META-ANALYSIS

Meta-analysis implies the aggregation of a field of data to summarize research findings and come to some general conclusions regarding effects or outcomes of a given treatment/project/program. There may be a standardization of measures taken from a range of tests or instruments to develop a manageable frame for an overall data analysis. Hence, the data from five or six experiments/projects/programs can be analyzed together to determine in some statistical manner what overall outcomes/effects were. Although meta-analysis is often conceptualized in terms of analyzing and scaling statistical data, it can also mean an in-depth analysis of sorts when examining the characteristic variables of a given treatment/project/program. Meta-analysis has been used in several fields of research.

Any in-depth systematic analysis of a set of characteristic variables

(findings) and summarization of findings after examining several sets of characteristic variables (findings) about a treatment/project/program can be called meta-analysis. Meta-analysis is based on research findings from completed research efforts.

Gene Glass, in a 1976 article, defines meta-analysis:

> Meta-analysis refers to the analysis of analyses. I use it to refer to the statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings. It connotes a rigorous alternative to the casual, narrative discussions of research studies which typify our attempts to make sense of the rapidly expanding research literature.

Meta-Analysis Technique

Glass' actual meta-analysis approach involves statistically transforming the findings of individual studies to a common metric by coding the various characteristics of the studies, then using conventional statistical

> ...procedures to determine whether there is an overall effect, subsample effects, and relations among characteristics of the studies and the findings. The original data for each unit of analysis in a study are not used. Rather, the unit of analysis is the study, and summary data from each study are analyzed.

What Glass did was to standardize various effect measures and control for these in analyzing data.

In meta-analysis a multivariate approach is most often preferred to control for systematic differences in the statistical analyses. The characteristics of the primary studies are, therefore, crucial to possible relations in conflicting or varying findings. Therefore, sampling is an important issue.

Although some researchers are rigorous about randomization, it is the nature of meta-analysis that if one wishes to analyze representative studies, one can depart from a randomized sample to a hand-picked sample that is most

representative.

Excluding primary studies for methodological design, weaknesses of sorts may not necessarily bias findings. This has been Glass' attitude.

Several statistical techniques can be used for data synthesis. See Tables I and II which follow and display techniques for data synthesis and meta-analytic calculations.

Meta-Analysis: Review of the Literature and Relevant Studies

A most comprehensive review of the literature pertaining to Meta-Analysis (Integrative Reviews) was done by the Washington Center for the Study of Services. This review points out that although it is quite common for federal and state governments or agencies to conduct integrative interviews--there is not necessarily a field of literature pertaining to methods and techniques pertaining to such reviews.

The literature referenced for this area stems back to 1971 (Feldman, and Light and Smith). In Light and Smith's 1971 article, they discuss how there is a "lack of systematic effort to accumulate information from a set of disparate studies." This lack of review methodologies hinders researchers from using the accumulation of previous research and the range of current research.

Gene Glass' 1976 article, "Primary, Secondary and Meta-Analysis of Research," further defined and conceptualized meta-analysis. Glass goes further here defining data analysis in terms of primary analysis, secondary analysis and meta-analysis.

An extensive meta-analysis of psychotherapy outcome studies was conducted by Mary Lee Smith and Gene V. Glass in 1977 when results of over 400 controlled evaluations of psychotherapy and counseling were coded and integrated statistically. Glass' conclusions after extensive systematic

statistical data analysis indicated that the results of research demonstrate counseling and psychotherapy have beneficial effects. This meta-analysis led to several critiques from psychologists and a rebuttal regarding the results of the study, the analysis techniques used, the quality of the studies (data) available to Glass for his research, and their underlying standards of judgment in conducting such an integrative review (Gallo and Eysenck 1978), (Glass 1978).

A major methodological criticism of Glass' psychotherapy meta-analyses was that he included a range of good, bad or indifferent reports and some of which were poorly designed. He believed that the sophistication and validity of the studies' designs were not so relevant when findings were the actual entity to be used and integrated with findings from other studies. The value of any meta-analysis research will always heavily rely on the validity and reliability of all of the underlying studies or data sets upon which it is based. Therefore, it is crucial that studies/findings/data selected for a meta-analysis are valid, reliable and sound in methodology.

The integration of studies with conflicting findings (Light and Smith 1971, and Light 1979) has been a theme running through the meta-analysis literature. Light advises researchers to use outcome variation as a "signal", rather than treating it as a "noise" in data synthesis efforts. Variations in conflicting findings can offer insights into the program's context. Reasons for conflicting findings on a given topic may be due to the fact that the phenomena investigated in the studies varied. More correlations of varying findings could be done to find systematic relationships in the end. The failure to do so results in inconclusive reviews or confusion around topics.

Since 1965 and the passage of the Elementary and Secondary Education

Act (ESEA) federal funds have been directed toward the special educational needs of disadvantaged children through compensatory education, bilingual education, etc. There have been several meta-analysis efforts at a national level to determine the impact of these categorically funded programs.

A cost effectiveness study "Tempo" was initiated in 1967 to examine Title I compensatory education programs nationally. Analysts found a lack of primary data--that the LEAs (local education agencies) had not collected the necessary information on program accomplishments and that the conceptual model of a cost-benefit analysis was not appropriate. They failed to show significant gains in student achievement.

In 1968 and 1969 Glass used a survey approach nationally to gather information on Title I programs. LEAs responded regarding descriptive information about programs but were defensive about achievement data. 1968 results showed no gain scores in student achievement and 1969 results were never made public. Out of 180,000 questionnaire mailed, only 8,000 to 10,000 were returned with usable student achievement data. A USOE in-house study estimated that at least twice as much achievement data were available than were reported by LEAs. The 1968 survey did suggest that Title I funds were in some cases used as general and not categorical aid for poor children.

After the 1968 and 1969 surveys, Title I case studies of exemplary progrmas were done by AIR through site visits and describing program populations, strategies and achievements of individual programs. This "It Works" case study approach resulted in a final report, "A Study of Selected Exemplary Programs for the Education of Disadvantaged Children."

In 1972 AIR was contracted by the Office of Program Planning and Evaluation (DHEW/OE) to reanalyze Title I data from FY 1965 through 1970. This extensive study by Michael Wargo and others reanalyzed national data

in terms of the following: Title I Management Performance; The ESEA Title I Context; The Needs of Children in Title I Schools; Resource Allocations (All levels); and The Impact of ESEA Title I on Participating Children.

Several data sources were used for this national study including interviews with federal personnel and annual reports. State Title I Annual Evaluation Reports were found to be the most comprehensive source for statewide and regional data.

A cost benefit type of study was included here with statistical evidence of the relationship between cost and effectiveness in reading projects.

Overall findings from this study indicated that most states failed to implement the Title I programs in full compliance, funds were misallocated to children without specific critical needs for the program and that aside from teacher opinion, there was little evidence at a national level that the program had a positive impact on participating children.

No evidence could be found at the state or local level that countered the noncompliance and resource allocation conclusions; however, positive impact was demonstrated by a few states, and nearly 100 local projects were identified that demonstrated significant cognitive benefits. Nevertheless, until the program and local projects are in full compliance with regulations, guidelines, and program criteria, it was felt that no conclusions could be made regarding the potential effectiveness of the compensatory educaton program intended by ESEA Title I.

In 1974 AIR conducted another national study, this time to determine the status of federally-funded bilingual education programs. This study was quite comprehensive and involved a meta-analysis of student achievement data while controlling many independent variables. A two-year evaluation of 38 Title VII projects in eight states was done for the Office of Education and

the reauthorization hearings of ESEA in 1978. Findings indicated that less than 1/3 of the students in Title VII programs were limited-English-speaking ability; that exit criteria for bilingual programs had not been established and that findings appeared to indicate that Hispanic students not in Title VII classes outperformed students in Title VII classes in English proficiency.

Although this report was extensive and controlled for the socioeconomic background of the students, language proficiencies of the students, teachers and staff, time in program, language usage in the classroom, instructional methodologies and other variables of the nature of the Title VII student population from 1975-77 hinders the usage of these findings in making general statements about the potential effects of bilingual education for non- or limited-English-speakers.

Recently(1980-82)there has been an interest in aggregating data on federally-funded programs as there have been substantial cut-backs in social programs under the Reagan Administration. Several contracts have been funded to evaluate the instructional component of bilingual education programs, to develop evaluation models and to identify effective programmatic variables. Title I programs are collecting data through developed evaluation models that were a result of federally-funded contracts to the RMC Research Corporation. There are currently new publications that address data aggregation issues.

Meta-analysis is of extreme value to policy and decision makers at a federal level. Since there is such a range of research studies on dozens of topics relevant to policymaking decisions, a summation of findings is essential as a basis for setting priorities for funding and program emphasis.

Since there is such a wide range of implementation strategies in Title I (Compensatory Education) and Title VII (Bilingual Education) (and often an

overlap of services), meta-analysis efforts have been somewhat inconclusive and refuted by professionals in related fields. Research efforts need to continue in this area.

II. META-EVALUATION

Meta-evaluation is the in-depth analysis of evaluation procedures, methodologies, activities, instruments used, interpretation of data, and general evaluation of information needed to determine the merit of a given treatment/program/project. Meta-evaluations are based on proposed evaluation intentions or findings as written in completed evaluation reports.

Meta-evaluation is based on evaluation designs or completed reports without normalizing data, scaling scores, etc. Hence, a meta-evaluation is a method of evaluation research clearly examining evaluation methodologies, procedures, data analysis techniques, interpretation of results and the validity and reliability of conclusions.

Meta-evaluation is a valuable research methodology in critiquing the value of proposed evaluation designs and completed evaluations. Meta-evaluation as a method of evaluation research has the potential of verifying or improving data collection efforts for accountability or program needs. At a local, state or federal level, data are aggregated to determine program effectiveness.

Meta-Evaluation: Review of the Literature and Relevant Studies

P.T. Orata wrote an article in 1940 called, "Evaluating Evaluation." M. Scriven in 1969 wrote "An Introduction to Meta-Evaluation." Here Scriven defined meta-evaluation, "Meta-evaluation is second-order evaluation. Theoretically, meta-evaluation involves the methodological assessment of the role of evaluation; practically, it is concerned with the evaluation of

specific performances." Since then, there have been other articles that address the systematic ways of estimating the worth of evaluations. Various checklists, formats, and standards have been proposed in evaluating the quality of evaluations.

Stake's "Table of Contents for a Final Evaluation Report" (1969) could serve as such a recommended format for determining the quality of an evaluation design or report. This "Table of Contents" focuses on"

I. Objectives of Evaluation
II. Specifications of the Program
III. Program Outcomes
IV. Relationships and Indicators
V. Judgments of Worth

A 1973 essay by B. Worthen, "A Look at the Mosaic of Educational Evaluation and Accountability," covered the following considerations in regards to evaluation quality.

1. Conceptual Clarity
2. Characteriztion of Program
3. Recognition and Representation of Legitimate Audiences
4. Sensitivity to Political Problems in Evaluation
5. Specification of Information Needs and Sources
6. Comprehensiveness/Inclusiveness
7. Technical Adequacy
8. Consideration of Project Cost
9. Explicit Standards/Criteria
10. Judgments and/or Recommendations
11. Reports Tailored to Audiences

D. Stufflebeam's 1974 "Meta-Evaluation" covers meta-evaluation, its purpose, and the process. In this article, he also proposed "An Administrative Checklist for Reviewing Evaluation Plans". This checklist points to a number to be considered under the:

I. Conceptualization of the Evaluation
II. Socio-political Factors
III. Contractual/Legal Arrangements between Evaluation and Client
IV. Technical Design
V. Management Plan for the Evaluation

Scriven (1974) proposed the use of his "Key Evaluation Checklist" and its possible use in meta-evaluation efforts. Items on this checklist for

evaluating products, produces, and proposals involves:

1. Need (Justification)
2. Market (Disseminability)
3. Performance (True Field Trial)
4. Performance (True Consumer)
5. Performance (Critical Comparison)
6. Performance (Long-Term)
7. Performance (Side Effects)
8. Performance (Process)
9. Performance (Causation)
10. Performance (Statistical Significance)
11. Performance (Educational Significance)
12. Costs and Cost-Effectiveness
13. Extended Support

A 1976 article by Sanders and Nafziger entitled "A Basis for Determining the Adequacy of Evaluation Designs" is also part of the meta-evaluation literature. The checklist includes the following areas of consideration for evaluation quality:

Scope	Reliability	Pervasiveness
Relevance	Objectivity	Ethical Considerations
Flexibility	Representativeness	Protocol
Feasibility	Timeliness	

(Exhibit II)

Cook and Gruder (1978) reviewed four projects aimed at evaluating the technical quality of summative evaluations implemented from 1971 to 1975. These meta-evaluations analyzed data from federally-funded projects and the Russell-Sage Foundation. Modes of meta-evaluation research were defined in this article (Table I).

Here meta-evaluation is presented in a different aspect in implying that the process can involve the manipulation of data--a simultaneous secondary analysis of raw data.

An April 1979 article entitled "QUEMAC VALUE: An Approach to Meta-Evaluation" was done by B. Gowan. He proposed critiquing evaluations by the following concepts:

Q = Question--Telling question of evaluation
U = Unquestioned Assumptions--By the workers of the evaluation
E = Events/ objects--Description of product
M = Method--Technical methods of evaluation
A = Answer--Claims, Conclusions
C = Concepts/ Conceptual Structure--Key Concepts

Cook and Gruder, 1978

TABLE I

MODELS OF META-EVLAUATION

Subsequent to Primary Evalaution	Data not manipulated	Single Data Sets	1. Essay review of an evaluation report.
		Multiple Data Sets	2. Review of the literature about a specific program.
	Data manipulated	Single Data Sets	3. Empirical reevaluation of an evaluation or program.
		Multiple Data Sets	4. Empirical reevaluation of multiple data sets about the same program.
Simultaneous with Primary Evaluation	Data not Manipulated	Single or Multiple Data Sets	5. Consultant meta-evaluation.
	Data Manipulated	Single Data Sets	6. Simultaneous secondary analysis of raw data.
		Multiple Data Sets	7. Multiple independent replications.

Meta-Evaluation and Bilingual Education

Regarding actual meta-evaluation studies, there is an overlap between what has been considered meta-analysis, secondary analysis, and meta-evaluation. There are frequently reviews of research and evaluation studies to consider the value or possibilities of the studies for making generalizations about given fields.

In the area of bilingual education, for example, in 1974 the Center for the Study of Evaluation conducted a project entitled "Evaluation and Decision Making: The Title VII Experience" (Alkin, Kosecoff, Fitz-Gibbon, Selegman, 1976). This project involved a meta-evaluation of 42 nationwide Title VII evaluation reports to determine their impact on the decisionmaking of both managers and sponsors. Results presented the characteristics of the evaluations, statements on quality and that there was a local use of the evaluation data, but that evaluations were not used as a basis for refunding at a sponsor's level. Quality issues here were based on:

- General Evaluation Designs
- Assessment Techniques
- Testing Instruments
- Interpretation of Results of Analysis
- Analysis of Raw Data
- Appropriateness of Data Collection Techniques
- Format of Evaluation Report
- Accuracy of Gathered Data
- Completeness of Report

In 1977, Zappert and Cruz did a review of over 180 studies to examine their designs and results. In 1979, Dulay and Burt examined 175 project evaluations.

In 1979, a meta-evaluation study was conducted by the California State Department of Education's Office of Bilingual-Bicultural Education. Twenty-five Title VII project proposals were reviewed over a three-year period. Results indicated evaluation designs as stated needed more specifics on intentions and desired outcomes.

The author's dissertation involved a meta-evaluation of 47 Title VII evaluation designs and reports in FY 1979-80. A 153-item checklist was used to rate designs and reports.

The following "CES Suggested Criteria for Bilingual Program Evaluation" suggest elements of meta-evaluation and evaluation designs which should be considered. Results of complete information coverage give more power to the evaluation for future meta-analyses studies and secondary analyses. Tables II, III, and IV show the format of the meta-evaluation instruments developed for the author's study. Many dimensions of the bilingual evaluations were considered.

The results of this study indicated that Title VII projects still need to improve their evaluation designs and reports before evaluation reports can truly be considered as data bases for future meta-analyses studies or as statements on the effectiveness of bilingual education.

P.Martin 7/80

CES META-EVALUATION CHECKLIST

Key Summary Sheet

Design
1 = Yes
_ = No

Degree Proposed
1. Partially Developed/Proposed
2. Developed/Proposed
3. Fully Developed/ Proposed
0. No Information/ NA

Degree Implemented
1. Partially Implemented
2. Implemented but could be Improved
3. Fully Implemented
0. No Information/ NA

Perceived Quality
1. Below Average
2. Average/Needs Improvement
3. Above Average
0. NA

ITEM	Degree Proposed	Degree Implemented	Perceived Quality
I. EVALUATION MODEL CHARACTERISTICS			
II. INFORMATION COVERAGE			
III. DESIGN CHARACTERISTICS (GENERAL)			
IV. DESIGN CHARACTERISTICS (SPECIFIC)			
V. DATA COLLECTION TECHNIQUES			
VI. TESTING SCHEDULE			
VII. TESTS			
VIII. ANALYTIC PROCEDURES			
IX EVALUATION UTILITY			

CES META-EVALUATION CHECKLIST

Project # _____	Design	Degree Proposed	Report	Degree Implemented	Perceived Quality	
III. DESIGN CHARACTERISTICS (GENERAL)						45-49
37. Purpose of Evaluation						50-54
38. Evaluation Objectives						55-59
39. Evaluation Timeline						60-64
40. Limitations/Constraints Discussed						65-69
41. Discrepancy of Intended and Observed						70-74
42. Cause and Effect Factors Addressed						75-79
43. Unanticipated Outcomes (General)						Dup. 1-2 Card No. 6(4) 5-9
44. Educational Significance						10-14
45. Performance (Long-Term)						15-19
46. Special Study						20-24
47. Longitudinal Study						25-29
48. Current Needs Assessment Data						30-34
49. Community Involvement in Evaluation						35-39
50.						40-44
51.						45-49
52.						50-54

Design: 1 = Yes _ = No Report: 1 = Yes _ = No

Degree Proposed: 1. Partially Developed/Proposed 2. Developed/Proposed 3. Fully Developed/Proposed 0. NA

Degree Implemented: 1. Partially Implemented 2. Implemented but could be Improved 3. Fully Implemented 0. NA

Perceived Quality: 1. Incomplete/Not Comprehensive 2. Somewhat Satisfactory 3. Well Done 0. NA

©1980 Paula Martin

TXu55-188

Project No. _____ (1-2)

CES SUPPLEMENTAL DATA SHEET

Scale:

1. Incomplete/Not Comprehensive
2. Somewhat Satisfactory
3. Well Done
0. N.A.

Card No. 15 (3-4)

1.	Proposal (Overall)			5
2.	Organization			6
3.	Identification of Needs			7
4.	Definition of Goals/Objectives/Intents			8
5.	Resources for Evaluation			9
6.	Clarity			10
7.	Readability			11
8.				12
9.				13
10.				14
11.	Evaluation Design (Overall)			15
12.	Current Needs Assessment Data			16
13.	Stated Purpose of Evaluation			17
14.	Definition of Audiences for Report/Information			18
15.	Timeline for Data Collection			19
16.	Description of Instruments			20
17.	Specification of Conditions - Testing, etc.			21
18.	Identification and Sampling of Participants/ Respondents			22
19.	Monitoring Program Activities to Objectives			23
20.	Overall Plan			24
21.	Flexibility			25
22.				26
23.				27
24.	Assessment Techniques (Overall)			28
25.	Testing schedule			29
26.	Sampling Techniques			30
27.	Control/Comparison Group Plan			31
28.	Testing Instruments to Information Needs			32
29.	Validity of CRTs			33
30.	Questionnaires			34
31.	Interview Formats			35
32.	Observations/Visitations/Reviews			36
33.	Other Documentation			37
34.	Data Storage and Retrieval			38
35.	Cost Analysis			39
36.	Overall Data Collection Procedures			40
37.				41
38.				42
39.				43

40.	Data Analysis (Overall)			44
41.	Statistical Tests			45
42.	Inferences			46
43.	Presentation of Data			47
44.	Technical Adequacy			48
45.	Conclusions (Interpretation of Results)			49
46.				50
47.				51
48.				52
49.	Report Characteristics (Overall)			53
50.	Validity			54
51.	Reliability			55
52.	Generalizability			56
53.	Credibility			57
54.	Objectivity			58
55.	Recommendations			59
56.	Utility of Evaluation Results (Overall)			60
57.	Timeliness			61
58.	Interim Reports			62
59.	Feedback to Staff/Respondents			63
60.	Several Audiences Addressed			64
61.	Program Planning			65
62.	Program Revisions			66
63.				67
64.				68
65.				69

Paula Martin
CES 7/80

Paula Martin, Ed.D.
Evaluation Consultant
CES 10/80

CES CRITERIA FOR THE EVALUATION OF BILINGUAL PROGRAMS

A complete evaluation design should consider information regarding the following:

I. Evaluation Model Characteristics
   1. Type of Model
      A. Descriptive
      B. Goal-based
      C. Objective-based
      D. Formative/Summative
      E. Specific theoretical model (e.g., CIPP)
   2. Funds for the Evaluation
   3. Evaluation Resources
   4. Staffing for the Evaluation

II. Information Coverage (Scope of the Report)
   1. Total Program
   2. Selected Components - Which?
   3. Student Achievement Data - Which areas, subjects?
   4. Cost information
   5. Definition of the Audiences for the Evaluation

III. General Design Characteristics
   1. Purpose of Evaluation
   2. Evaluation Objectives
   3. Evaluation Timeline
   4. Discussion of Limitations/Constraints
   5. Discrepancy of intended objectives and observed performance
   6. Discussion of cause and effect factors
   7. Allowance for noting unanticipated outcomes
   8. Current Needs Assessment Data
   9. Possible special studies and longitudinal studies

VI. Specific Design Characteristics

1. Testing schedule and population included
2. Tests to be used - Standardized - Criterion Referenced
3. Norms Used for Reference
4. Control or Comparison Groups
5. Sampling Techniques
6. Entry/Exit Criteria
7. Other standards, levels of minimum competency, continuums of skills, etc.

V. Data Collection Types and Techniques

1. Questionnaires - To whom?
2. Interviews (Telephone or Face-to-Face) - To whom?
3. Reports from Project Staff
4. Special Data Collection
   (a) suspensions/disciplinary records/dropout rate
   (b) attendance patterns (students)
   (c) attendance patterns, logs, calendars of activities and other records
5. Special Evaluation Activities
   (a) use of evaluation teams
   (b) coordination with State or Federal program review efforts
   (c) use of parents/community in evaluation

VI. Analytic Procedures

1. Type of Statistics
   (a) means
   (b) standard deviations
   (c) percentiles
   (d) T-test
   (e) regression techniques
2. Special Approaches
   (a) longitudinal analysis
   (b) special comparison studies/groups
3. Data Processing Techniques
   (a) SPSS
   (b) BIOMED

4. Quantified Ratings from Observations/Questionnaires
5. Documentation and No Statistical Procedures
6. Summaries and narrative reports
7. Other Analysis
8. Validity and Reliability Issues

VII. Evaluation Utility

1. Schedule of reports and type of reports
2. Audiences for evaluation information
3. Dissemination of evaluation results
4. Use of results for planning, decision-making, program information.

III. SECONDARY ANALYSIS

Secondary analysis as defined by Glass is: "The re-analysis of data for the purpose of answering the original research question with better statistical techniques or answering new questions with old data."

As a research methodology, secondary analysis can be considered an important resource for educational research and evaluation (Burstein, 1978) because "it has the potential to consider important questions without some of the limitations, or with a different set of limitations, than those encountered in the original investigation." Positive aspects of secondary analysis according to Burstein include:

1. Different perspectives of data for different conclusions.
2. Psychological distance from the original data.
3. Provides substantive and methodological "anchor to the past and speedier routes" to the future.
4. Secondary analysts do not have the time constraints, psychological constraints, or economic constraints that the original researchers may have faced.
5. Secondary analysis can be a fundamental in training educational researchers.

Regarding secondary analysis, to ensure that access to data is less of a problem, certain considerations regarding data usage and archives must be considered. Hedrick, Boruch and Ross, 1977, proposed:

> (1) Original investigators should anticipate the likelihood of secondary analysis.
> (2) Secondary analysis should be bound by the same confidentiality and privacy restrictions as the primary analysts and should be held legally responsible for any breach of ethics in the conduct of the secondary analysis.

Primary concerns of the secondary analyst involve:

1. Access to data
2. Sampling appropriate studies from a population of primary studies
3. Reliability and validity of the study's original design
4. Lost data sets
5. Time constraints
6. Confidentiality of data

Sampling is an important issue in secondary analyses where primary data is screened for its appropriateness in an integrated secondary analysis.

There have been some significant secondary analyses including: the Motseller-Moynihan secondary analysis of the Colemen study; the Campbell-Erlebacher analysis of the Ohio-Westinghouse Headstart evaluation; and the Elashoff-Snow secondary analysis of Pygmalion in the Classroom.

As Bryant and Wortman (1978) state in their article "Secondary Analysis: The Case for Data Archives," it is becoming increasingly important to preserve research data so that it can be reexamined and reanalyzed by others.

Conclusions

Meta-analysis and meta-evaluation are relevant to the field of education. Secondary analysis has special considerations in maintaining files with original data for re-analysis.

At a national level, state level and school district level, information has to be summarized on program effectiveness. Concrete data on student achievement is needed to convince Congress, state legislatures, and local school boards, as well as parents, that educational institutions and special programs are doing what they intend to do.

Needless to say, it is not an easy task to summarize any findings in the social science areas (concerning treatments/project/programs) since there is a range of context variables, implementation characteristics, and resources available. Projects also use different measuring devices, collect different types of data. Educational programs have many confounding variables and it is difficult if not impossible to isolate the effect of any one program on overall outcomes in terms of student achievement. Programs are impacted by several funding sources and within-school differences.

Since meta-analysis efforts involve a statistical summation of a collection of data analyses, results from individual program evaluations need to follow a more standardized reporting format and approach toward analyzing data, and presenting student achievement results for state and national data collection purposes. Evaluation reports for state and federally-funded programs could be improved to become actual data bases for meta-analysis studies. Reliable evaluation findings could be aggregated to draw some general conclusions regarding the effects of special programs, particular instructional approaches, methodologies, materials in use, and other aspects.

Since evaluations are conducted as part of the state and federal regulations, meta-evaluation should be built into the process to continue to improve current data collection efforts.

The value of any meta-analysis research relies heavily on the validity and reliability of all the underlying studies or data sets upon which it is based. This make meta-evaluation important as a first step toward successful meta-analyses for local, state, or federal information purposes where a summation of findings regarding special programs is essential as a basis for policymaking decisions regarding priorities for program emphasis, funding, or legislation. Meta-evaluation also can continue to point out the difficulty of isolating the effects of any one project treatment or variable

as it relates to program outcomes or student achievement. This can aid in the development of better overall data collection processes that would present reliable data on program effects.

As current budgetary cut-backs affect state and federal programs, meta-analysis is becoming an important issue in documenting program effects. Meta-avaluation as well should play a more important role in improving data collection efforts.

Meta-Evaluation References

Alkin, M., Kosecoff, J., Fitz-Gibbon, C., & Seligman, R. Evaluation and decision making: The Title VII experience. CSE Monograph Service in Evaluation, No. 4. Los Angeles: University of California, Center for the Study of Evaluation, 1976.

Boruch, R. F., Cordray, D. S., Pion, G., & Leviton, L. A mandated appraisal of evaluation practices: Digest of recommendations to the Congress and Department of Education. Contract #300-79-0467 for the U.S. Department of Education. June 30, 1980.

Rentz, C. C. and R. R. Rentz, eds. Evaluating federally sponsored programs. San Francisco: Jossey-Bass Publishers, No. 2, 1978.

Sanders, J. R. & Nafziger, D. H. A basis for determining the adequacy of evaluation designs. Occasional Paper Series, No. 6. Kalamazoo, Mich.: Western Michigan University, The Evaluation Center, 1976.

Scriven, M. Evaluation bias and its control. Occasional Paper Series, No. 4. Kalamazoo, Mich.: Western Michigan University, The Evaluation Center, 1975.

_____. Evaluation thesaurus, 2nd ed. Pt. Reyes, Calif.: Edgepress, 1980.

_____. An introduction to meta-evaluation. In P. A. Taylor & D. M. Cowley, eds., Readings in curriculum evaluation. Dubuque, Iowa: Wm. C. Brown Co. Publishers, 1972. (a)

_____. Key evaluation checklist. Pt. Reyes, Calif.: Edgepress 1974.

_____. The methodology of evaluation. In P. A. Taylor & D. M. Cowley, eds., Readings in curriculum evaluation. Dubuque, Iowa: Wm. C. Brown Co. Publishers, 1972. (b)

Sheppard, Lorrie, "A Checklist for Evaluating Large Scale Assessment Programs." Occasional Paper Series, No. 9. Kalamazoo, Mich.: Western Michigan University.

Smith, N., ed. Federal efforts for developing new evaluative methods. San Francisco: Jossey-Bass Publishers, 1981.

Spolsky, B., ed. Advances in language testing. Rosslyn, Va.: Center for Applied Linguistics, 1975-78.

Stake, R. E. Evaluation design, instrumentation, data collection, and analysis of data. In J. L. David, ed., Educational evaluation. Columbus, Ohio: State Stuperintendent of Public Instruction, 1969. Reprinted with revisions in B. R. Worthen & J. R. Sanders, eds., Educational evaluation: Theory and practice. Worthington, Ohio: Charles A. Jones, 1973.

Additional References

Meta-Analysis

Campbell, D. T. & Boruch, R. Making the case for randomized assignment to treatments by considering the alternatives: Six ways in which quasi-experimental evaluations in compensating education tend to underestimate effects. In C. A. Bennett & A. A. Lumsdaine (Eds.), Evaluation and experiment. New York: Academic Press, 1975.

Cicirelli, V. G., et al. The Impact of Head Start: An Evaluation of the Effects of Head Start on Children's Cognitive and Affective Development. Washington, D.C.: Office of Economic Opportunity, 1969.

Cohen, D. K. The Politics of evaluation. Review of Educational Research, 1970, 40, 270.

Coleman, J. S., et al. Equality of educational opportunity. Washington, D.C.: U.S. Department of Health, Education and Welfare, U.S. Office of Education, OE-38001, 1966, 218-333.

Cook, T. D. The potential and limitations of secondary evaluation. In M. S. Apple, M. S. Subkoviak, & J. R. Lutler (Eds.), Educational evaluation analysis and responsibility. Berkeley, CA: McCutchan, 1974.

Elashoff, J. D. & Snow, R. E. (Eds.). Pygmalion reconsidered. Worthington, Ohio: Charles A. Jones, 1971.

Gilbert, J. P., Light, R. J., & Mosteller, F. Assessing social innovations: An empirical base for policy. In C. A. Bennett & A. A. Lumsdaine (Eds.). Evaluation and experiment. New York: Academic Press, 1975.

Glass, G. V. Integrating findings: The meta-analysis of research. Review of Research in Education, 1977, 351-379.

Hedrick, T. E., Boruch, R. F., & Ross, J. Policy and regulation for ensuring the availability of evaluation data for secondary analysis. Paper presented at the Annual Meeting of the American Psychological Association, San Francisco, September 1977.

Levin, H. M. A new model of school effectiveness. Do teachers make a difference. Washington, D.C.: U.S. Department of Health, Education and Welfare, Office of Education, Bureau of Educational Personnel Development (OE-58042), 1970, 55-775.

Light, R. J. & Smith, P. V. Accumulating evidence: Procedures for resolving contradictions among different studies. Harvard Educational Review, 1971, 41, 429-271.

Pillemer, D. B. & Light, R. J. Using the results of controlled experiments to construct social programs: Three caveats. Evaluation studies review annual. Beverly Hills, Calif.: Sage Publishing Company, 1979.

Powell, M. Necessary steps to insure availability of data for secondary analysis. Paper presented at the Annual Meeting of the American Educational Research Association, New York, April 1977.

Rosenthal, R. Combining results of independent studies. Psychological Bulletin, 1978, 85, 185-193.

Rosenthal, R. & Jacobson, L. Pygmalion in the classroom: Teacher expectation and pupils' intellectual development. New York: Holt, Rinehart & Winston, 1968.

Simpson, E. H. Interpretation of interaction in contingency tables. Journal of the Royal Statistical Society, Series B, 1951, 13(2), 238-241.

Smith, M. S. & Bissell, J. S. Report analysis: The impact of Head Start. Harvard Education Review, 1970, 40, 51-104.

Wiley, D. E. & Harnischfeger, A. Explosion of a myth: Quantity of schooling and exposure to instruction, major educational vehicles. Studies of Educative Processes, 1974, No. 8. University of Chicago. (Also: Educational Researcher, 1974, 3, 7-12.)

Stufflebeam, D. L. Meta-evaluation. Occasional Paper Series, No. 3. Kalamazoo, Mich.: Western Michigan University, The Evaluation Center, 1974.

Summary of a study of DSEA Title VII proposals, 1975-1978. Sacramento: California State Department of Education, Office of Bilingual-Bicultural Education.

Wooldridge, Ronald J., ed. Evaluation of complex systems. San Francisco: Jossey-Bass Publishers, No. 10, 1981.

An evaluation of product information packages (PIPs) as used for the diffusion of bilingual projects. (RMC Report UR460.) Mountain View, Ca.: RMC Research Corporation, 1980.

Martin, Paula H., Evaluation in bilingual education programs; a meta-evaluation of Title VII projects, FY 1979-80. Dissertation study for the University of San Francisco: San Francisco, CA, 1981.

References

Meta-Analysis, Secondary Analysis, and Content Analysis

Ball, Samuel, ed. Assessing and interpreting outcomes. San Francisco: Jossey-Bass Publishers, No. 9, 1981.

Boruch, R. F., ed. Secondary analysis. San Francisco: Jossey-Bass Publishers, 1978.

Bryant, F. B. & Wortmen, P. M. Secondary analysis: The case for data archives. American Psychologist, April 1978.

Burstein, L. Secondary analysis: An important resource for educational research and evaluation. Educational Researcher, May 1978.

ESEA, Title I: A reanalysis and synthesis of evaluation data from fiscal year 1965 through 1970. (AIR-27400-3/72). Palo Alto: American Institute for Research.

Evaluation of the classroom instruction component of the ESEA Title VII Bilingual Education Program. Development Associates, Inc. in affiliation with Abt Associates, Inc. for the Office of Program Evaluation, U.S. Department of Education.

Evaluation of the impact of ESEA Title VII Spanish/English bilingual education programs: Overview of study and findings. (AIR-48300-3/78-FR VI). Palo Alto: American Institute for Research, 1978. (ERIC Report ED 154-6340).

Eysenck. "An Exercise in Mega-Silliness," American Psychologist, May 1978, pp. 517-518.

Gallo, Phillip S. "Meta-Analysis--A Mixed Meta-phor," American Psychologist, May 1978, pp. 515-517.

Glass, G. V. Policy for the unpredictable (uncertain research and policy). Educational Researcher, October 1979.

_____. Primary, secondary and meta-analysis of research. Educational Researcher, December 1976, pp. 3-8.

_____. Standards and criteria. Occasional Paper Series, No. 10. Kalamazoo, Mich.: Western Michigan University, The Evaluation Center, 1977.

_____, McGaw, B., & Smith, M. L. Meta-analysis in social research. Beverly Hills, Calif.: Sage Publications, 1981.

Horst, D. P., Tallmadge, G. K., & Wood, C. T. A practical guide to measuring project impact on student achievement. No. 1 in a series of monographs on evaluation in education. Washington, D.C.:

Government Printing Office, No. 017-080-0146-2. Prepared for HEW by RMC Research Corporation, 1975.

Krippendorff, K. Content analysis. Beverly Hills, Calif.: Sage publications, 1980.

Light, R. J. Capitalizing on variation: How conflicting research findings can be helpful for policy. Educational Researcher, 1979, 9(9).

_____. Research integration: The state of the art. Evaluation in Education, 1980, 4(1).

Roberts, K. & L. Burstein, eds. Issues in Data Aggregation, No. 6. San Francisco: Jossey-Bass Publishers, 1980.

Scriven, Michael. Evaluation thesaurus, 2nd ed. Inverness, Calif.: Edge Press, 1980.

Smith, M. L. & Glass, G. V. Meta-analysis of psychotherapy outcome studies. American Psychologist, 1977, 32, 752-760.

Tallmadge, G. K. & Horst, D. P. A procedural guide for validation of achievement gains in educational projects. No. 2 in a series of monographs on evaluation in education. Washington, D.C.: Government Printing Office, No. 017-080-01516. Prepared for HEW by RMC Research Corporation, 1971.

Wargo, M. J., et al. ESEA Title I: A re-analysis and synthesis of evaluation data from fiscal year 1965 through 1970: Final report. (AIR-27400-3/72 FR). Palo Alto, Calif.: American Institute for Research, 1972.